Preliminary Data Requirements for IOC-2

D. Beckman, B. Bush, K. Nagel,
P. Medvick, R. Smith, P.Stretz, M. Williams

Los Alamos National Laboratory

22 July 1997

Abstract

The TRANSIMS case study in Portland will require a variety of population, activity, network, land use, and meteorological data. This presentation provides an overview of the data requirements for this case study. It outlines general data collection issues, enumerates the types of data needed and their attributes, and discusses possibilities for automated data generation and validation.

Outline

- *overview*
- *general questions*
- *data preferences*
- *specific data*
  - *network data*
  - *population and activity data*
  - *transit data*
  - *freight data*
  - *air-quality data*
  - *measurements for calibration and validation*
- *generating missing data*
- *validating data sets*
- *proposal for data collection*

Overview

- *IOC-1* focused on the microsimulation research and development for a single-mode case study.*
  - *Detailed network data formats are available for IOC-1.*
  - *No formal population, activity, or planning data specifications were developed for IOC-1.*
- *IOC-2 will focus on population, activity, and planning research and development for a multiple-mode case study.*
  - *The network data format specification will be revised to include the description of new modes.*
  - *Formal population, activity, and planning data specifications will be developed.*
- *Caveats:*
  - *Research is still ongoing, so not all of our data needs are clearly defined and final specifications are not available yet.*
  - *The case study has not been chosen yet.*

* IOC = interim operational capability

General Questions

- *Do you have comprehensive data catalogs that are easy for us to peruse?*
- *How difficult will it be to obtain data for local streets throughout the greater Portland region?*
- *Are you aware of methods . . .*
  - *to develop algorithms for generating missing data?*
  - *to develop tools for validating the consistency of data?*
- *How long will it take to collect the various types of data mentioned later in this presentation?*
- *Can you provide data formatted according to TRANSIMS specifications (when such specifications are available)?*
- *Is high-resolution aerial photography available in digital format?*
- *What is the general quality of the TIGER/Line data in the Portland area?*

Data Preferences

- *The date (i.e., the year) for the network data should match the date used for population, transit, and other data. The age of data should be noted.*
- *Indications of data quality such as accuracy, completeness, and original source are valuable.*
- *The preferred measurement units are SI (meters, seconds, meters per second, etc.).*
- *The preferred geographic coordinate system is UTM (NAD 27).*
- *The preferred delivery formats are delimited text files, dBASE tables, ArcView shape files, and Arc/Info coverages.*

Network Data

- *Elements:*
  - *nodes*
  - *links*
  - *pocket lanes*
  - *lane uses*
  - *lane connectivity*
  - *unsignalized intersections*
  - *signalized intersections*
  - *parking locations*
  - *mode transfer locations*
  - *mode crossing points*
- *We also need to know how network elements vary with time-of-day (e.g., phasing plans for traffic lights).*
- *It would be useful to know the TIGER/Line record id numbers corresponding to the various network elements.*

Nodes

- *A node is a junction between links.*
- *High-priority attributes:*
  - *id number*
  - *x coordinate*
  - *y coordinate*

Links

- *A link connects two nodes.*
- *They may be streets, rail lines, bikeways, etc.*
- *They are unidirectional.*
- *High-priority attributes:*
  - *id number*
  - *name (e.g. street name)*
  - *node id numbers for endpoints*
  - *number of through lanes*
  - *length (must not be shorter than straight-line map distance)*
  - *grade*
  - *capacity*
  - *speed limit*
  - *functional class (i.e, RTP designation)*
  - *modes of travel and their means of separation*
  - *toll*

Links (continued)

- *Low-priority attributes:*
  - *“setbacks” at intersections*
  - *“free-flow” speed*
  - *“crawl” speed*
  - *default “through” link at ending node*
  - *id number of reverse link*
- *Estimates can be used for low-priority data attributes.*

Pocket Lanes

- *Pocket lanes are non-through lanes on links.*
- *Low-priority attributes:*
  - *link id number*
  - *lane number*
  - *position along link*
  - *type (merge pocket, turn pocket, or pull-out), including on freeways*
  - *length*
- *For parts of the network not critical to the case study, it might be possible to generate this data automatically, or omit it altogether.*

Lane Uses

- *Any special lane uses (HOV lanes, restrictions on trucks, etc.) must be noted.*
- *High-priority attributes:*
  - *link number id*
  - *lane number*
  - *type of use*

Lane Connectivity

- *Lane connectivity specifies how lanes are connected across a node.*
- *High-priority attributes:*
  - *prohibited turns*
- *Low-priority attributes:*
  - *node id number*
  - *table of which incoming lanes connect to which outgoing lanes*
  - *turn penalties*
- *We have algorithms to generate this data automatically where it is not available.*

Unsignalized Intersections

- *Unsignalized intersections have sign controls (as opposed to having traffic lights).*
- *All nodes are either unsignalized or signalized.*
- *Low-priority attributes:*
  - *node id number*
  - *the sign on each incoming link at each intersection (stop, yield, or none)*
- *Given general rules for placing signs at intersections, it is possible to generate this data automatically where it is not available, provided at least one "through" direction is given.*

Signalized Intersections

- *Signalized intersections have a traffic light.*
- *High-priority attributes:*
  - *node id number*
  - *type of signal (pre-timed)*
- *Medium-priority attributes:*
  - *phasing plan (which movements are allowed in each phase and how they are protected)*
  - *timing plan (how long the phases last and how they are sequenced)*
  - *timing offsets*
- *We are interested in developing algorithms that can generate the medium-priority attributes from the characteristics of the incoming links, such as impedances or capacities.*
- *We might develop microsimulation techniques that do not require the low-priority attributes to be specified at all.*

Parking Locations

- *Parking areas are located along links.*
- *They may represent actual driveways or parking places, or be generic.*
- *Low-priority attributes:*
  - *id number*
  - *link id number*
  - *position along link*
  - *type (parallel on street, head-in on street, driveway, parking lot entrance, or generic)*
  - *capacity*
  - *pricing*
- *For regions of the network not critical to the case study, it will be possible to generate this data automatically.*
- *Major centers of production or attraction may require extra detail.*

Mode Transfer Points

- *Transfer points are where travelers change travel modes.*
- *High-priority attributes:*
  - *id number*
  - *beginning and ending modes*
  - *beginning and ending link numbers*
- *Low-priority attributes:*
  - *name*
  - *positions along beginning and ending links*
  - *time delays (including walking times in stations)*
  - *capacity*
  - *pricing*
- *Estimates can be used for low-priority data attributes.*

Mode Crossing Points

- *Mode crossings are where a link of one mode crosses (and interferes with) a link of a different mode, but where there is no opportunity for travelers to change their mode of travel.*
- *High-priority attributes:*
  - *link id numbers for the modes*
  - *positions along links where the crossing occurs*
  - *type of interference*
  - *time delays*
  - *right-of-way*

Other Network Elements

- *Depending on the case study chosen, several other network data elements might be modeled in IOC-2:*
  - *roadway traffic sensors*
  - *actuated traffic signals*
  - *wide-area traffic controls*
  - *ITS technology*
  - *anything that significantly slows down dynamics of cars, trains, bicycles, pedestrians, etc. (e.g., speed bumps, pedestrian crossings)*

Population and Activity Data

- *map of traffic analysis zones*
- *population location and demographics*
- *activities and surveys of activities*
- *tours*
- *mode choice information*
- *trip tables*
  - *household trips*
  - *taxis*
  - *freight*
- *land use*
  - *attractors (e.g., local businesses*)*
  - *producers*

* Would commercially-available business directories be useful here?

Population and Activity Data (continued)

- *vehicle information*
  - *registered street address*
  - *vehicle age and type*
- *non-vehicular modes (i.e., bicycle and pedestrian)*
- *any special Census Bureau or other data sets not publicly available*

Transit Data

- *Elements:*
  - *routes*
  - *schedules*
  - *vehicles*
- *We have not yet determined at what resolution/fidelity we will simulate transit.*

Transit Routes

- *A transit route is the path a transit vehicle takes in the network.*
- *High-priority attributes:*
  - *id number*
  - *name*
  - *mode*
  - *sequence of link id numbers*

Transit Schedules

- *A transit schedule tells when an individual transit vehicle arrives at each point on its route.*
- *High-priority attributes:*
  - *route number*
  - *stop locations*
  - *headways*
  - *dwell times*
  - *layovers*
  - *vehicle*
  - *pricing*
  - *time control points*
- *We also need information on operational rules and procedures (e.g., what do buses do when the get behind schedule or fill to capacity).*
- *Riderships surveys would be useful, if available.*

Transit Vehicles

- *A transit vehicle carries travelers along a transit route.*
- *High-priority attributes:*
  - *type of vehicle (bus, train, etc.)*
  - *capacity (sitting and standing)*
  - *age*
  - *technology*

Freight Data

- *O-D matrices for freight*
- *representation in trip tables*
- *freight routes*
- *roads with freight restrictions*
- *heavy freight vs. light commercial*
- *truck fleet composition*

Air Quality Data

- *emission inventories for point sources and non-transportation activities*
- *upper-air data (probably from the airport)*
- *previous short-term air quality studies*
- *inputs to airshed models*
  - *land use*
  - *emissions*
- *sources of data (e.g., State of Oregon, Department of Environmental Quality, Air Quality Division, EPA, University researchers, and the National Climatic Data Center?)*

Measurements for Calibration and Validation

- *turn counts*
- *traffic volumes*
- *chase car and instrumented vehicle studies*
- *road sensor data*
- *transit performance (e.g., delays from schedule)*
- *intersection capacities or impedances*
- *air monitoring data*
- *videotape*
- *other*

Generating Missing Data

- *It is possible to generate automatically data that is too expensive or difficult to collect.*
- *In IOC-1, we estimated several types of data:*
  - *parking locations*
  - *“through” links at intersections*
  - *lane connectivity*
  - *signs at unsignalized intersections*
- *Additional capabilities would be useful for IOC-2:*
  - *local street properties*
  - *road grades*
  - *pocket lane locations*
  - *intersection geometry*
  - *signalized intersection locations, phasing plans, timing plans, and timing offsets*
  - *mode transfer locations*

Validating Data Sets

- *The TRANSIMS planner and microsimulation require consistent and realistic input data.*
- *Planning networks generally are not adequate for microsimulation. (IOC-1 required enhancements to freeways, local streets, and major production/attraction areas.)*
- *We have already developed procedures for identifying data problems:*
  - *data field values*
  - *cross-references between id numbers in files*
  - *network topology (connectivity, dead-ends, etc.)*
  - *lane connectivity*
  - *allowed movements at intersection controls*
  - *phasing and timing plans for signals*
  - *lengths vs. Euclidean distances*
- *Automated methods of checking IOC-2 data also need to be developed.*

Proposal for Data Collection

1. *Determine the availability of the data mentioned in this presentation.*
2. *Obtain raw data (i.e., data in any usable format) of interest for research purposes.*
3. *Continue LANL population, activity, planner, and microsimulation research for IOC-2.*
4. *Team up with organizations to develop techniques for generating missing data and for validating data sets.*
5. *Finalize IOC-2 data format specifications.*
6. *Receive final data for case study in IOC-2 format.*
7. *Perform Portland case study.*